

PAGE: 1

RAW SEQUENCE LISTING
PATENT APPLICATION US/09/394,519

DATE: 09/24/1999
TIME: 14:33:38

Input Set: I394519.RAW

This Raw Listing contains the General Information
Section and up to first 5 pages.

ENTERED

1 <110> APPLICANT: Heard, Jacqueline
2 Riechmann, Jose Luis
3 Adam, Luc
4 Broun, Pierre
5 Pineda, Omaira
6 Reuber, Lynne
7 Jiang, Cai-Zhong
8 Keddie, James
9 Zhang, James
10 Benito, Maria-Ines
11 Yu, Guo-Liang
12 Fromm, Mike
13 <120> TITLE OF INVENTION: PLANT GENE SEQUENCES I
14 <130> FILE REFERENCE: MBI-0003
15 <140> CURRENT APPLICATION NUMBER: US/09/394,519
16 <141> CURRENT FILING DATE: 1999-09-13
17 <150> EARLIER APPLICATION NUMBER: 60/101,349
18 <151> EARLIER FILING DATE: 1998-09-22
19 <150> EARLIER APPLICATION NUMBER: 60/103,312
20 <151> EARLIER FILING DATE: 1998-10-06
21 <150> EARLIER APPLICATION NUMBER: 60/108,734
22 <151> EARLIER FILING DATE: 1998-11-17
23 <150> EARLIER APPLICATION NUMBER: 60/113,409
24 <151> EARLIER FILING DATE: 1998-12-22
25 <160> NUMBER OF SEQ ID NOS: 170
26 <170> SOFTWARE: PatentIn Ver. 2.0
27 <210> SEQ ID NO 1
28 <211> LENGTH: 1577
29 <212> TYPE: DNA
30 <213> ORGANISM: Arabidopsis thaliana
31 <220> FEATURE:
32 <221> NAME/KEY: CDS
33 <222> LOCATION: (90)..(1214)
34 <223> OTHER INFORMATION: G4
35 <400> SEQUENCE: 1
36 aaagaatcga atatttatta ttctgccccg aagattctat ttctgatcat ttacaccocct 60
37 aaaaagagta gagctttcgt gaagccacc atg tgt gga gga gct ata atc tcc 113
38 Met Cys Gly Gly Ala Ile Ile Ser
39 1 5
40 gat ttc ata cct ccg ccg agg tcc ctc cgc gtc act aac gag ttt atc 161
41 Asp Phe Ile Pro Pro Pro Arg Ser Leu Arg Val Thr Asn Glu Phe Ile
42 10 15 20
43 tgg ccg gat ctg aaa aac aaa gtg aaa gct tca aag aag aga tcg aat 209
44 Trp Pro Asp Leu Lys Asn Lys Val Lys Ala Ser Lys Lys Arg Ser Asn

PAGE: 2

RAW SEQUENCE LISTING
PATENT APPLICATION US/09/394,519

DATE: 09/24/1999

TIME: 14:33:38

Input Set: I394519.RAW

45	25	30	35	40	
46	aag cga tcc gat ttc ttc gat ctt gac gat gat ttc gaa gct gat ttc				257
47	Lys Arg Ser Asp Phe Phe Asp Leu Asp Asp Asp Phe Glu Ala Asp Phe				
48		45	50	55	
49	caa ggg ttt aag gat gac tcg gct ttt gac tgc gaa gac gat gat gat				305
50	Gln Gly Phe Lys Asp Asp Ser Ala Phe Asp Cys Glu Asp Asp Asp Asp				
51		60	65	70	
52	gtc ttc gtc aat gtt aag cct ttc gtc ttc acc gca act act aag ccc				353
53	Val Phe Val Asn Val Lys Pro Phe Val Phe Thr Ala Thr Thr Lys Pro				
54		75	80	85	
55	gta gct tcc gct ttc gtc tcc act gta ggt tca gca tat gcc aag aaa				401
56	Val Ala Ser Ala Phe Val Ser Thr Val Gly Ser Ala Tyr Ala Lys Lys				
57		90	95	100	
58	act gta gag tcc gct gag caa gct gag aaa tct tct aag agg aag agg				449
59	Thr Val Glu Ser Ala Glu Gln Ala Glu Lys Ser Ser Lys Arg Lys Arg				
60	105	110	115	120	
61	aag aat car tac cga ggg att agg cag cgt cct tgg gga aaa tgg gct				497
62	Lys Asn Gln Tyr Arg Gly Ile Arg Gln Arg Pro Trp Gly Lys Trp Ala				
63		125	130	135	
64	gcg gag atc cgt gat ccg aga aaa ggc tcc cga gaa tgg ctt gga aca				545
65	Ala Glu Ile Arg Asp Pro Arg Lys Gly Ser Arg Glu Trp Leu Gly Thr				
66		140	145	150	
67	ttc gac act gct gag gaa gca gca aga gct tat gat gct gca gca cgc				593
68	Phe Asp Thr Ala Glu Glu Ala Ala Arg Ala Tyr Asp Ala Ala Ala Arg				
69		155	160	165	
70	aga atc cgt ggc acg aaa gct aag gtg aat ttt ccc gag gag aag aac				641
71	Arg Ile Arg Gly Thr Lys Ala Lys Val Asn Phe Pro Glu Glu Lys Asn				
72		170	175	180	
73	cct agc gtc gta tcc cag aaa cgt cct agt gct aag act aat aat ctt				689
74	Pro Ser Val Val Ser Gln Lys Arg Pro Ser Ala Lys Thr Asn Asn Leu				
75	185	190	195	200	
76	cag aaa tca gtg gct aaa cca aac aaa agc gta act ttg gtt cag cag				737
77	Gln Lys Ser Val Ala Lys Pro Asn Lys Ser Val Thr Leu Val Gln Gln				
78		205	210	215	
79	cca aca cat ctg agt cag cag tac tgc aac aac tcc ttt gac aac tct				785
80	Pro Thr His Leu Ser Gln Gln Tyr Cys Asn Asn Ser Phe Asp Asn Ser				
81		220	225	230	
82	ttt ggt gat atg agt ttc atg gaa gag aag cct cag atg tac aac aat				833
83	Phe Gly Asp Met Ser Phe Met Glu Glu Lys Pro Gln Met Tyr Asn Asn				
84		235	240	245	
85	cag ttt ggg tta aca aac tcg ttc gat gct gga ggt aac aat gga tac				881
86	Gln Phe Gly Leu Thr Asn Ser Phe Asp Ala Gly Gly Asn Asn Gly Tyr				
87		250	255	260	
88	cag tat ttc agt tcc gat cag ggc agt aac tcc ttc gac tgt tct gag				929
89	Gln Tyr Phe Ser Ser Asp Gln Gly Ser Asn Ser Phe Asp Cys Ser Glu				
90	265	270	275	280	
91	ttc ggg tgg agt gat cac ggc cct aaa aca ccc gag atc tct tca atg				977
92	Phe Gly Trp Ser Asp His Gly Pro Lys Thr Pro Glu Ile Ser Ser Met				
93		285	290	295	
94	ctt gtc aat aac aac gaa gca tca ttt gtt gaa gaa acc aat gca gcc				1025

PAGE: 3

RAW SEQUENCE LISTING PATENT APPLICATION US/09/394,519

DATE: 09/24/1999
TIME: 14:33:38

Input Set: I394519.RAW

```

95      Leu Val Asn Asn Asn Glu Ala Ser Phe Val Glu Glu Thr Asn Ala Ala
96              300              305              310
97      aag aag ctc aaa cca aac tct gat gag tca gac gat ctg atg gca tac      1073
98      Lys Lys Leu Lys Pro Asn Ser Asp Glu Ser Asp Asp Leu Met Ala Tyr
99              315              320              325
100     ctt gac aac gcc ttg tgg gac acc cca cta gaa gtg gaa gcc atg ctt      1121
101     Leu Asp Asn Ala Leu Trp Asp Thr Pro Leu Glu Val Glu Ala Met Leu
102             330             335             340
103     ggc gca gat gct ggt gct gtg act cag gaa gag gaa aac cca gtg gag      1169
104     Gly Ala Asp Ala Gly Ala Val Thr Gln Glu Glu Glu Asn Pro Val Glu
105             345             350             355             360
106     cta tgg agc tta gat gag atc aat ttc atg ctg gaa gga gac ttt      1214
107     Leu Trp Ser Leu Asp Glu Ile Asn Phe Met Leu Glu Gly Asp Phe
108             365             370             375
109     tgaagtgatc gatggttcct tagtttgtaa ataaagctgt gttggatttt gctgttgggg      1274
110     gatggtacaa gtcacacctc aagctctatg cattggtatc tcatgagcct tctcttccat      1334
111     agagagtttc tcttttaatt ttgtcgaaat aaaaaagggtg tgatgaagta aatagaggta      1394
112     taataatata tctctattaa gtcttgtttt gttctttcat ttttgatttt cttttctatt      1454
113     taaaagacag tttattagtc ttctgagctc tctttttgat ctttggtata gcgtatcatc      1514
114     accctcgaaa gtgtaatgtt ttgtaccccc aaacttgttt agcattataa taaagtctct      1574
115     ttg      1577
116 <210> SEQ ID NO 2
117 <211> LENGTH: 375
118 <212> TYPE: PRT
119 <213> ORGANISM: Arabidopsis thaliana
120 <220> FEATURE:
121 <223> OTHER INFORMATION: G4
122 <400> SEQUENCE: 2
123     Met Cys Gly Gly Ala Ile Ile Ser Asp Phe Ile Pro Pro Pro Arg Ser
124             1             5             10             15
125     Leu Arg Val Thr Asn Glu Phe Ile Trp Pro Asp Leu Lys Asn Lys Val
126             20             25             30
127     Lys Ala Ser Lys Lys Arg Ser Asn Lys Arg Ser Asp Phe Phe Asp Leu
128             35             40             45
129     Asp Asp Asp Phe Glu Ala Asp Phe Gln Gly Phe Lys Asp Asp Ser Ala
130             50             55             60
131     Phe Asp Cys Glu Asp Asp Asp Asp Val Phe Val Asn Val Lys Pro Phe
132             65             70             75             80
133     Val Phe Thr Ala Thr Thr Lys Pro Val Ala Ser Ala Phe Val Ser Thr
134             85             90             95
135     Val Gly Ser Ala Tyr Ala Lys Lys Thr Val Glu Ser Ala Glu Gln Ala
136             100            105            110
137     Glu Lys Ser Ser Lys Arg Lys Arg Lys Asn Gln Tyr Arg Gly Ile Arg
138             115            120            125
139     Gln Arg Pro Trp Gly Lys Trp Ala Ala Glu Ile Arg Asp Pro Arg Lys
140             130            135            140
141     Gly Ser Arg Glu Trp Leu Gly Thr Phe Asp Thr Ala Glu Glu Ala Ala
142             145            150            155            160
143     Arg Ala Tyr Asp Ala Ala Ala Arg Arg Ile Arg Gly Thr Lys Ala Lys
144             165            170            175

```

PAGE: 4

RAW SEQUENCE LISTING
PATENT APPLICATION US/09/394,519

DATE: 09/24/1999
TIME: 14:33:38

Input Set: I394519.RAW

```

145      Val Asn Phe Pro Glu Glu Lys Asn Pro Ser Val Val Ser Gln Lys Arg
146                180                      185                      190
147      Pro Ser Ala Lys Thr Asn Asn Leu Gln Lys Ser Val Ala Lys Pro Asn
148                195                      200                      205
149      Lys Ser Val Thr Leu Val Gln Gln Pro Thr His Leu Ser Gln Gln Tyr
150                210                      215                      220
151      Cys Asn Asn Ser Phe Asp Asn Ser Phe Gly Asp Met Ser Phe Met Glu
152      225                230                      235                      240
153      Glu Lys Pro Gln Met Tyr Asn Asn Gln Phe Gly Leu Thr Asn Ser Phe
154                245                      250                      255
155      Asp Ala Gly Gly Asn Asn Gly Tyr Gln Tyr Phe Ser Ser Asp Gln Gly
156                260                      265                      270
157      Ser Asn Ser Phe Asp Cys Ser Glu Phe Gly Trp Ser Asp His Gly Pro
158                275                      280                      285
159      Lys Thr Pro Glu Ile Ser Ser Met Leu Val Asn Asn Asn Glu Ala Ser
160                290                      295                      300
161      Phe Val Glu Glu Thr Asn Ala Ala Lys Lys Leu Lys Pro Asn Ser Asp
162      305                310                      315                      320
163      Glu Ser Asp Asp Leu Met Ala Tyr Leu Asp Asn Ala Leu Trp Asp Thr
164                325                      330                      335
165      Pro Leu Glu Val Glu Ala Met Leu Gly Ala Asp Ala Gly Ala Val Thr
166                340                      345                      350
167      Gln Glu Glu Glu Asn Pro Val Glu Leu Trp Ser Leu Asp Glu Ile Asn
168                355                      360                      365
169      Phe Met Leu Glu Gly Asp Phe
170                370                      375
171      <210> SEQ ID NO 3
172      <211> LENGTH: 1699
173      <212> TYPE: DNA
174      <213> ORGANISM: Arabidopsis thaliana
175      <220> FEATURE:
176      <221> NAME/KEY: CDS
177      <222> LOCATION: (417)..(1418)
178      <223> OTHER INFORMATION: G5
179      <400> SEQUENCE: 3
180      tttttttttt gcaatctccc cctaattctgt tgtttctcgc ttctttcttct gttaatcatc 60
181      tgtctttcaa aaagaaagaa aaaagaaaaa ttcgatttct gggtttgttt ttgtcataca 120
182      gaaaaaaatc aagcttatga atttgtgttt aattttttgt tttaatttga aaggcagggt 180
183      ttttcagaac gagatcgttt tttcaaattt cttctgattt tacctctttt tttcttctta 240
184      gatttttagtg aatcgagggt gaaatttttg attcctcttt ttcggatcta cacagagggt 300
185      gcttatttca aaccttttag atccattttt ttttaatttt ctcggaaaaa tccctgtttc 360
186      tttacttttt tataagtctc aggttcaatt ttttcggatt caaattttta ttttaa atg 419
187                                     Met
188                                     1
189      gca gct gct atg aat ttg tac act tgt agc aga tcg ttt caa gac tct 467
190      Ala Ala Ala Met Asn Leu Tyr Thr Cys Ser Arg Ser Phe Gln Asp Ser
191                5                      10                      15
192      ggt ggt gaa ctc atg gac gcg ctt gta cct ttt atc aaa agc gtt tcc 515
193      Gly Gly Glu Leu Met Asp Ala Leu Val Pro Phe Ile Lys Ser Val Ser
194                20                      25                      30

```

PAGE: 5

RAW SEQUENCE LISTING
PATENT APPLICATION US/09/394,519

 DATE: 09/24/1999
 TIME: 14:33:38

Input Set: I394519.RAW

195	gat tct cct tct tct tct tct gca gcg tct gcg tct gcg ttt ctt cac	563
196	Asp Ser Pro Ser Ser Ser Ser Ala Ala Ser Ala Ser Ala Phe Leu His	
197	35 40 45	
198	ccc tct gcg ttt tct ctc cct cct ctc ccc ggt tat tac ccg gat tca	611
199	Pro Ser Ala Phe Ser Leu Pro Pro Leu Pro Gly Tyr Tyr Pro Asp Ser	
200	50 55 60 65	
201	acg ttc ttg acc caa ccg ttt tca tac ggg tcg gat ctt caa caa acc	659
202	Thr Phe Leu Thr Gln Pro Phe Ser Tyr Gly Ser Asp Leu Gln Gln Thr	
203	70 75 80	
204	ggg tca tta atc gga ctc aac aac ctc tct tct tct cag atc cac cag	707
205	Gly Ser Leu Ile Gly Leu Asn Asn Leu Ser Ser Ser Gln Ile His Gln	
206	85 90 95	
207	atc cag tct cag atc cat cat cct ctt cct ccg acg cat cac aac aac	755
208	Ile Gln Ser Gln Ile His His Pro Leu Pro Pro Thr His His Asn Asn	
209	100 105 110	
210	aac aac tct ttc tcg aat ctt ctc agc cca aag ccg tta ctg atg aag	803
211	Asn Asn Ser Phe Ser Asn Leu Leu Ser Pro Lys Pro Leu Leu Met Lys	
212	115 120 125	
213	caa tct gga gtc gct gga tct tgt ttc gct tac ggt tca ggt gtt cct	851
214	Gln Ser Gly Val Ala Gly Ser Cys Phe Ala Tyr Gly Ser Gly Val Pro	
215	130 135 140 145	
216	tcg aag ccg acg aag ctt tac aga ggt gtg agg caa cgt cac tgg gga	899
217	Ser Lys Pro Thr Lys Leu Tyr Arg Gly Val Arg Gln Arg His Trp Gly	
218	150 155 160	
219	aaa tgg gtg gct gag atc cgt ttg ccg aga aat cgg act cgt ctc tgg	947
220	Lys Trp Val Ala Glu Ile Arg Leu Pro Arg Asn Arg Thr Arg Leu Trp	
221	165 170 175	
222	ctt ggg act ttt gac acg gcg gag gaa gct gcg ttg gcc tat gat aag	995
223	Leu Gly Thr Phe Asp Thr Ala Glu Glu Ala Ala Leu Ala Tyr Asp Lys	
224	180 185 190	
225	gcg gcg tac aag ctg cgc ggc gat ttc gcc cgg ctt aac ttc cct aac	1043
226	Ala Ala Tyr Lys Leu Arg Gly Asp Phe Ala Arg Leu Asn Phe Pro Asn	
227	195 200 205	
228	cta cgt cat aac gga ttt cac atc gga ggc gat ttc ggt gaa tat aaa	1091
229	Leu Arg His Asn Gly Phe His Ile Gly Gly Asp Phe Gly Glu Tyr Lys	
230	210 215 220 225	
231	cct ctt cac tcc tca gtc gac gct aag ctt gaa gct att tgt aaa agc	1139
232	Pro Leu His Ser Ser Val Asp Ala Lys Leu Glu Ala Ile Cys Lys Ser	
233	230 235 240	
234	atg gcg gag act cag aaa cag gac aaa tcg acg aaa tca tcg aag aaa	1187
235	Met Ala Glu Thr Gln Lys Gln Asp Lys Ser Thr Lys Ser Ser Lys Lys	
236	245 250 255	
237	cgt gag aag aag gtt tcg tcg cca gat cta tcg gag aaa gtg aag gcg	1235
238	Arg Glu Lys Lys Val Ser Ser Pro Asp Leu Ser Glu Lys Val Lys Ala	
239	260 265 270	
240	gag gag aat tcg gtt tcg atc ggt gga tct cca ccg gtg acg gag ttt	1283
241	Glu Glu Asn Ser Val Ser Ile Gly Gly Ser Pro Pro Val Thr Glu Phe	
242	275 280 285	
243	gaa gag tcc acc gct gga tct tcg ccg ttg tcg gac ttg acg ttc gct	1331
	Glu Glu Ser Thr Ala Gly Ser Ser Pro Leu Ser Asp Leu Thr Phe Ala	

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Input Set: I394519.RAW

Line	Error/Warning	Original Text
405	W Invalid/Missing Amino Acid Numbering	
642	W Invalid/Missing Amino Acid Numbering	
1685	W "N" or "Xaa" used: Feature required	tatattaaaa ggttggttaga tatananana nnnnnnnn
1706	W "N" or "Xaa" used: Feature required	Ala Ser Ser Thr Met Ser Ser Ser Thr Xaa S
1750	W "N" or "Xaa" used: Feature required	cct tct tct act gca agt ggt aat gat att a
1807	W "N" or "Xaa" used: Feature required	Thr Ala Ser Gly Asn Asp Ile Asn Thr Leu M
1881	W Invalid/Missing Amino Acid Numbering	
1884	W "N" or "Xaa" used: Feature required	atcatatgct atgttatgta tagtattata aaaaaatg
2013	W "N" or "Xaa" used: Feature required	tctgaataat caatgtcttt gcccttttn naangntn
2372	W "N" or "Xaa" used: Feature required	tgacataatg gttttgatgt tgtggagaag atagagag
2374	W "N" or "Xaa" used: Feature required	gtctttgaaa agcctttaga tgttttntaa ttcgtaag
2617	W Invalid/Missing Amino Acid Numbering	
3333	W "N" or "Xaa" used: Feature required	ttctttgggt ctctggtggt gggaggacta agnagtac
3565	W "N" or "Xaa" used: Feature required	agttangaaa cattgtaatt tgatttccaa attaattct
4070	W "N" or "Xaa" used: Feature required	agtttgcagc atgaaagctg attaaatatc agtcccgt
4121	W "N" or "Xaa" used: Feature required	attctctctc tcncaaaatt tatttctctc ggcttctt
4591	W "N" or "Xaa" used: Feature required	caaaaaaaaaat ctctcccatt aaaagnctgc ccaaagaa
5287	W "N" or "Xaa" used: Feature required	tataatccnt tgtgatactg aaaaaaaaaa aaaaaaaa
5401	W "N" or "Xaa" used: Feature required	aaaaagggtc acccccattt tatcttatgn nnaaagan
5646	W Invalid/Missing Amino Acid Numbering	
8137	W Invalid/Missing Amino Acid Numbering	
9738	W Invalid/Missing Amino Acid Numbering	
10483	W "N" or "Xaa" used: Feature required	ttagaaatcc acaatatctt tcttcacact ttntttca
10543	W Invalid/Missing Amino Acid Numbering	